



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2018 – 05:43 pm GMT

PDB ID : 4AX8  
Title : Medium resolution structure of the bifunctional kinase- methyltransferase WbdD  
Authors : Hagelueken, G.; Huang, H.; Naismith, J.H.  
Deposited on : 2012-06-11  
Resolution : 3.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

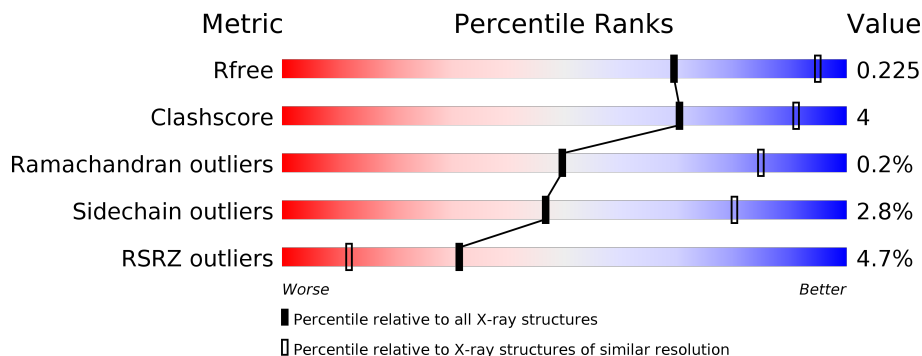
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	111664	1851 (3.00-3.00)
Clashscore	122126	2167 (3.00-3.00)
Ramachandran outliers	120053	2101 (3.00-3.00)
Sidechain outliers	120020	2104 (3.00-3.00)
RSRZ outliers	108989	1751 (3.00-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	569	<div> <div>4%</div> <div> <div></div> <div>70%</div> <div>7%</div> <div>21%</div> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7361 atoms, of which 3611 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

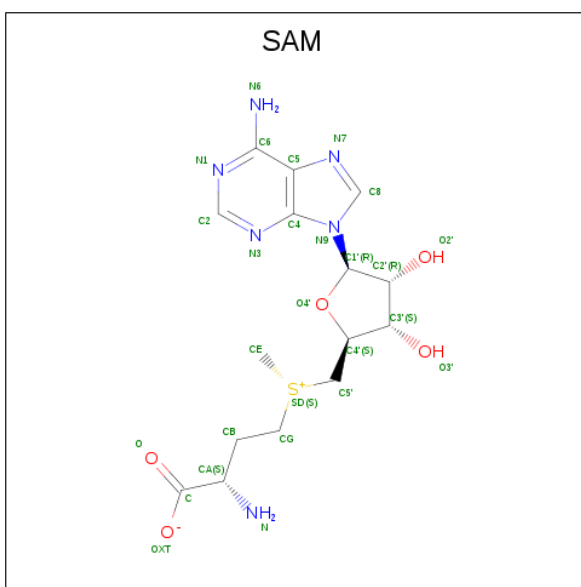
- Molecule 1 is a protein called WBDD.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	450	7262	2373	3577	629	673	10	0	6	0

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-12	MET	-	expression tag	UNP Q47592
A	-11	HIS	-	expression tag	UNP Q47592
A	-10	HIS	-	expression tag	UNP Q47592
A	-9	HIS	-	expression tag	UNP Q47592
A	-8	HIS	-	expression tag	UNP Q47592
A	-7	HIS	-	expression tag	UNP Q47592
A	-6	HIS	-	expression tag	UNP Q47592
A	-5	GLU	-	expression tag	UNP Q47592
A	-4	ASN	-	expression tag	UNP Q47592
A	-3	LEU	-	expression tag	UNP Q47592
A	-2	TYR	-	expression tag	UNP Q47592
A	-1	PHE	-	expression tag	UNP Q47592
A	0	GLN	-	expression tag	UNP Q47592
A	1	GLY	-	expression tag	UNP Q47592
A	168	PHE	LEU	conflict	UNP Q47592
A	273	TYR	HIS	conflict	UNP Q47592
A	440	VAL	ALA	conflict	UNP Q47592
A	480	VAL	GLY	conflict	UNP Q47592
A	518	GLN	PRO	conflict	UNP Q47592

- Molecule 2 is S-ADENOSYLMETHIONINE (three-letter code: SAM) (formula: C<sub>15</sub>H<sub>22</sub>N<sub>6</sub>O<sub>5</sub>S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	S	0	0
			49	15	22	6	5	1		

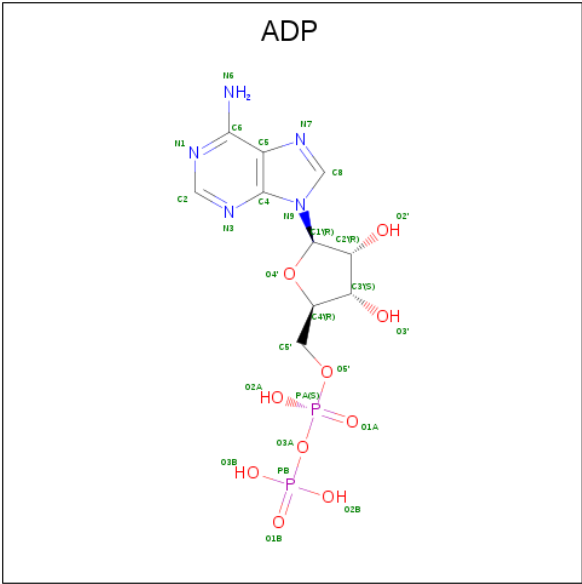
- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 5	O 4	S 1	0	0
3	A	1	Total 5	O 4	S 1	0	0

- Molecule 4 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula:

C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>10</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	P	0	0
			39	10	12	5	10	2		

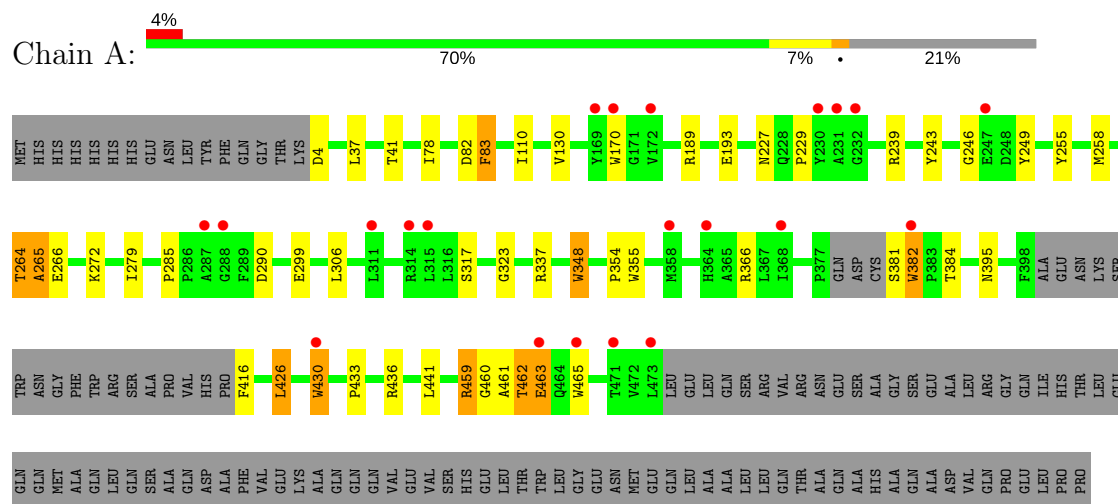
- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Cl	0	0
			1	1		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

#### • Molecule 1: WBDD



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	167.34Å 167.34Å 167.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	68.32 – 3.00 68.32 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.6 (68.32-3.00) 99.6 (68.32-3.00)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.93 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.212 , 0.233 0.204 , 0.225	Depositor DCC
$R_{free}$ test set	784 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	83.7	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 40.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.034 for -l,-k,-h	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7361	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.77% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SAM, SO4, ADP, CL

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.97	10/3803 (0.3%)	0.96	6/5169 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	265	ALA	N-CA	-7.22	1.31	1.46
1	A	430	TRP	CD2-CE2	7.20	1.50	1.41
1	A	382	TRP	CD2-CE2	6.01	1.48	1.41
1	A	348	TRP	CD2-CE2	5.90	1.48	1.41
1	A	465	TRP	CD2-CE2	5.70	1.48	1.41
1	A	285	PRO	C-O	-5.61	1.12	1.23
1	A	170	TRP	CD2-CE2	5.54	1.48	1.41
1	A	264	THR	CA-C	-5.54	1.38	1.52
1	A	285	PRO	CA-C	5.24	1.63	1.52
1	A	355	TRP	CD2-CE2	5.04	1.47	1.41

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	264	THR	CA-C-N	-12.39	89.95	117.20
1	A	264	THR	CA-C-O	7.54	135.93	120.10
1	A	189	ARG	NE-CZ-NH2	-6.75	116.93	120.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	189	ARG	NE-CZ-NH1	5.63	123.12	120.30
1	A	83	PHE	CB-CG-CD2	-5.31	117.08	120.80
1	A	264	THR	C-N-CA	-5.00	109.19	121.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	460	GLY	Peptide

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3685	3577	3576	27	0
2	A	27	22	22	2	0
3	A	10	0	0	1	0
4	A	27	12	12	3	0
5	A	1	0	0	0	0
All	All	3750	3611	3610	29	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

All (29) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:290:ASP:O	1:A:366:ARG:NH2	1.91	1.02
1:A:243:TYR:OH	4:A:1477:ADP:O1B	1.94	0.85
1:A:264:THR:O	1:A:265:ALA:C	2.30	0.65
1:A:229:PRO:O	4:A:1477:ADP:O3'	2.13	0.61
1:A:426:LEU:HD12	1:A:430:TRP:HD1	1.68	0.58
1:A:426:LEU:HD12	1:A:430:TRP:CD1	2.38	0.57
1:A:37:LEU:O	1:A:41:THR:HG23	2.10	0.52
4:A:1477:ADP:O1A	4:A:1477:ADP:O3B	2.28	0.51

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:78:ILE:HD12	1:A:78:ILE:C	2.31	0.50
1:A:426:LEU:CD1	1:A:430:TRP:CD1	2.95	0.49
1:A:290:ASP:C	1:A:366:ARG:NH2	2.61	0.49
1:A:246:GLY:HA3	1:A:249:TYR:CE2	2.50	0.47
1:A:382:TRP:HA	1:A:384:THR:H	1.79	0.47
1:A:459:ARG:O	1:A:461:ALA:HB2	2.16	0.46
1:A:348:TRP:CE3	1:A:382:TRP:CZ3	3.04	0.46
2:A:1474:SAM:HB1	2:A:1474:SAM:H4'	1.97	0.45
1:A:227:ASN:ND2	1:A:239:ARG:HD3	2.31	0.45
1:A:110:ILE:CD1	1:A:130:VAL:HG21	2.47	0.45
1:A:264:THR:C	1:A:266:GLU:N	2.65	0.45
1:A:272[B]:LYS:CD	1:A:299:GLU:OE1	2.66	0.43
1:A:110:ILE:HD11	1:A:130:VAL:HG21	2.00	0.43
1:A:317:SER:HB3	1:A:354:PRO:HB2	1.99	0.43
1:A:255:TYR:CB	1:A:258:MET:HE2	2.49	0.42
1:A:462:THR:O	1:A:463:GLU:C	2.58	0.42
1:A:82:ASP:OD2	2:A:1474:SAM:O3'	2.33	0.41
1:A:348:TRP:CZ3	1:A:382:TRP:HZ3	2.39	0.41
1:A:279:ILE:HG12	1:A:306:LEU:HD13	2.02	0.41
1:A:337:ARG:NE	3:A:1476:SO4:O1	2.39	0.40
1:A:433:PRO:HG2	1:A:436:ARG:HD3	2.02	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	450/569 (79%)	436 (97%)	13 (3%)	1 (0%)	49 85

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	323	GLY

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	392/486 (81%)	381 (97%)	11 (3%)	47 80

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	ASP
1	A	83	PHE
1	A	193	GLU
1	A	381	SER
1	A	395	ASN
1	A	416	PHE
1	A	426	LEU
1	A	441	LEU
1	A	459	ARG
1	A	462	THR
1	A	463	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 1 is monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SAM	A	1474	-	21,29,29	1.91	3 (14%)	16,42,42	2.44	4 (25%)
3	SO4	A	1475	-	4,4,4	0.54	0	6,6,6	0.89	0
3	SO4	A	1476	-	4,4,4	0.51	0	6,6,6	0.84	0
4	ADP	A	1477	-	25,29,29	1.35	3 (12%)	25,45,45	1.95	5 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SAM	A	1474	-	-	0/8/33/33	0/3/3/3
3	SO4	A	1475	-	-	0/0/0/0	0/0/0/0
3	SO4	A	1476	-	-	0/0/0/0	0/0/0/0
4	ADP	A	1477	-	-	0/12/32/32	0/3/3/3

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1474	SAM	CG-SD	-6.50	1.66	1.80
4	A	1477	ADP	C8-N9	-3.10	1.33	1.36
2	A	1474	SAM	C5-C4	2.30	1.45	1.40
4	A	1477	ADP	C5-C4	2.56	1.46	1.40
2	A	1474	SAM	O4'-C1'	3.31	1.45	1.41
4	A	1477	ADP	PB-O3A	3.56	1.65	1.60

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1474	SAM	N3-C2-N1	-7.33	122.59	128.86
4	A	1477	ADP	N3-C2-N1	-6.12	123.63	128.86
4	A	1477	ADP	PA-O3A-PB	-3.37	121.29	132.63
4	A	1477	ADP	O5'-C5'-C4'	-2.65	99.79	109.00
2	A	1474	SAM	O3'-C3'-C4'	-2.36	104.23	111.06
4	A	1477	ADP	O2A-PA-O1A	2.20	123.31	112.14
4	A	1477	ADP	O3B-PB-O2B	2.74	118.44	107.59
2	A	1474	SAM	C2'-C3'-C4'	3.03	108.43	102.62
2	A	1474	SAM	C4'-O4'-C1'	3.11	113.07	109.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1474	SAM	2	0
3	A	1476	SO4	1	0
4	A	1477	ADP	3	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	450/569 (79%)	0.49	21 (4%) 31 12	38, 66, 112, 139	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	169	TYR	5.4
1	A	231	ALA	5.1
1	A	471	THR	4.2
1	A	364[A]	HIS	4.1
1	A	170	TRP	3.8
1	A	465	TRP	3.6
1	A	230	TYR	3.4
1	A	368	ILE	3.3
1	A	172	VAL	3.0
1	A	463	GLU	2.5
1	A	358	MET	2.4
1	A	315	LEU	2.4
1	A	232	GLY	2.3
1	A	287	ALA	2.3
1	A	247	GLU	2.3
1	A	314	ARG	2.3
1	A	288	GLY	2.2
1	A	473	LEU	2.2
1	A	430	TRP	2.1
1	A	382	TRP	2.1
1	A	311	LEU	2.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	SO4	A	1476	5/5	0.74	0.27	111,115,126,129	0
5	CL	A	1478	1/1	0.85	0.18	78,78,78,78	0
4	ADP	A	1477	27/27	0.88	0.24	59,79,109,116	0
2	SAM	A	1474	27/27	0.96	0.23	49,60,68,74	0
3	SO4	A	1475	5/5	0.97	0.25	39,41,45,53	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.